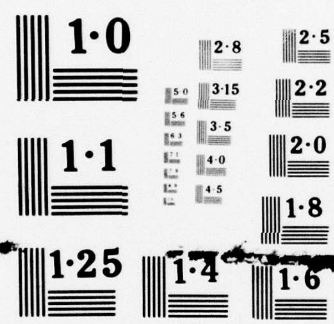
AD-A046 045 HARVARD UNIV CAMBRIDGE MASS DIV OF ENGINEERING AND --ETC F/G 12/2
OPTIMAL LINEAR SYSTEMS WITH POLYNOMIAL PERFORMANCE MEASURES,(U)
1973 R W BROCKETT N00014-75-C-0648

UNCLASSIFIED NL

1 OF 1
ADA
046045

END
DATE
FILMED
11-77
DDC




NATIONAL BUREAU OF STANDARDS
MICROCOPY RESOLUTION TEST CHART

AD A046045

WP16 - 1:30

(14)

(6)

# OPTIMAL LINEAR SYSTEMS WITH POLYNOMIAL PERFORMANCE MEASURES



(11) 1973

(12) 6p.

(10)

ROGER W. BROCKETT
Division of Engineering and Applied Physics
Harvard University
and
Center for Mathematical System Theory
University of Florida

Abstract

The linear time invariant aspects of the least squares theory is extended in this paper to linear time invariant systems with performance measures of the form

$$\eta = \int_0^\infty L(x) + ||u||^{2r} dt; \quad L(x) \geq 0; \quad L(0) = 0$$

where L is a polynomial in the components of x. Those L's which have optimal control laws which are linear are characterized, and the inverse problem, i.e. the characterization of those controllers which are optimal relative to a performance measure of this type is also solved. Our basic tools in studying this problem are a new lemma on the path independence of certain integrals which generalizes the earlier result on quadratic forms and a positivity condition investigated earlier by Jacques Willems and the author.

## 1. Introduction

R.E. Kalman's 1963 paper [1] characterizing, for linear stationary systems, those linear control laws which minimize time independent nonnegative quadratic performance measures on an infinite time interval, attracted considerable attention because the conditions were given in terms of the Nyquist locus and confirmed, to some extent, engineering practice. The original paper discussed the scalar input problem but it was shown subsequently that the multivariable problem is essentially the same. Other papers in this vein have appeared subsequently but always staying within the linear-quadratic framework.

In this paper we work out the corresponding theory for linear systems

$$\dot{x}(t) = Ax(t) + bu(t) \tag{1}$$

with performance measures of the rather more general form

$$\eta = \int_0^\infty L(x) + u^{2r} dt; \quad L \in P^+ \tag{2}$$

where $P^+$ denotes the set of polynomials which are nonnegative for all real arguments and vanish when all arguments do. As might be expected this theory leads to different conclusions regarding optimality, in general relaxing the conditions found for the quadratic case. Somewhat surprisingly we find that the L's which yield linear optimal control laws are necessarily homogeneous of degree 2r. Even among these, however, linear optimal control laws are the exception and not the rule provided r exceeds one.

A particular feature of Kalman's original result was that it was stated in such a way as to imply, together with the circle criterion, a strong stability property with respect to time varying and nonlinear gains. We find that even in our more general context this feature is preserved.

In this paper we use for the most part, the notation D = d/dt and write (1), (2) in higher order form as

$$p(D)y = u\ ; \tag{3}$$

and

$$\eta = \int_0^\infty L(y, y^{(1)}, \dots y^{(n-1)}) + u^{2r} dt; \quad L \in P^+ \tag{4}$$

In the last section we make the connection with the first order vector differential equation notation.

## 2. Path Independence

In our earlier work [2-6] we showed that the connection between frequency domain conditions and the existance of quadratic Liapunov functions and the connection between frequency domain conditions and quadratic optimality could be made using a lemma about the independence of path of integrals of the form

$$\eta = \int_a^b \sum \alpha_{ij} x^{(i)} x^{(j)} dt; \quad \frac{d^i x}{dt^i} = x^{(i)} \tag{5}$$

This lemma, which we will generalize below, enables one to avoid the use of Fourier transforms and manipulations of the Kalman-Yacubovic-Popov type in treating these problems. Moreover, using this approach we were in [5], able to bring the frequency domain theory to a new level of completion by avoiding positive definiteness conditions and hence allowing one to establish for the first time, frequency domain instability conditions. In this

(15)

This work was supported in part by the U.S. Office of Naval Research under the Joint Services Electronics Program Contract N00014-75-C-0648 and the Army Research Office under Grant DAAG29-76-C-0139





163 750

paper we take advantage of the intrinsic simplicity of this approach in using it as a point of departure to develop the theory of higher order problems of the form (3), (4).

The lemma of this section depends on certain elementary facts about the group of all permutations of a finite set of n symbols -- the symmetric group on n letters, S(n). Of course S(n) has n! elements; we denote a typical element by $\pi$. If we have an rth degree form in the variables $x_1, x_2, \ldots, x_n$ we can represent it by

$$f = \sum_{(a,b,\ldots c)\in I(n,r)} \alpha_{a,b,\ldots c} x_a x_b, \ldots, x_c \tag{6}$$

where $I(n,r) = \{(a,b,\ldots,c):\ 0\leq a\leq n,\ 0\leq b\leq n, \ldots 0\leq c\leq n;$ r terms$\}$. But such a representation is not unique unless we symmetrize. (This is a generalization of the symmetrization of the matrices representing quadratic forms.) A little thought will convince the reader that f may also be expressed as

$$f = \frac{1}{n!} \sum_{\pi\in S(n)} \sum_{(a,b,\ldots c)\in I(n,r)} \alpha_{\pi(a),\pi(b),\ldots,\pi(c)} x_a x_b \ldots x_c \tag{7}$$

In such a representation the $\alpha$'s corresponding to a specific (unordered) collection of subscripts on the x's are symmetric in that the interchange of two subscripts in $\alpha_{a,b,\ldots c}$ does not change its value.

In what follows we use $x^{(i)}$ to denote the ith derivative of the function $x : \mathbb{R}^1 \to \mathbb{R}^1$.

Lemma 1: If $x : \mathbb{R}^1 \to \mathbb{R}^1$ is n times differentiable then the integral

$$\eta = \int_a^b \sum_{(a,b,\ldots c)\in I(n,r)} \alpha_{a,b,\ldots c} \underbrace{x^{(a)} x^{(b)} \ldots x^{(c)}}_{\text{r factors}} dt \tag{8}$$

is expressible in terms of $x(a)$, $x^{(1)}(a), \ldots x^{(n-1)}(a)$ and $x(b)$, $x^{(1)}(b), \ldots x^{(n-1)}(b)$ if and only if the polynomial in the indeterminates $s_1, s_2, \ldots s_r$

$$h(s_1,s_2,\ldots s_r) = \frac{1}{n!} \sum_{\pi\in S(n)} \sum_{(a,b,\ldots c)\in I(n,r)} \alpha_{\pi(a),\pi(b),\ldots \pi(c)} s_1^a s_2^b \ldots s_r^c \tag{9}$$

can be factored as $h(s_1,s_2,\ldots s_r) = (s_1+s_2+\ldots s_r)\times m(s_1,s_2,\ldots s_r)$ with m a polynomial.

Proof: Without loss of generality we may suppose that the coefficients $\alpha$ have been symmetrized via (7). This does not change the value of the integral and means that we can express h more simply as

$$h(s_1,s_2,\ldots,s_r) = \sum_{(a,b,\ldots c)\in I(n,r)} \alpha_{a,b,\ldots c} s_1^a s_2^b \ldots s_r^c \tag{10}$$

Now suppose that the integral is independent of path in the above sense and that

$$\eta = v(x, x^{(1)}, \ldots, x^{(n-1)})\Big|_a^b = \sum_{(a,b,\ldots c)\in I(n-1,r)} \beta_{a,b,\ldots c} x^{(a)} x^{(b)} \ldots x^{(c)}\Big|_a^b \tag{11}$$

with the $\beta$'s being symmetrized. This being the case, the derivative of $\eta$ -- the integrand of the original expression for $\eta$ -- is just

$$\dot\eta = \sum_{(a,b,\ldots c)\in I(n-1,r)} \beta_{a,b,\ldots c} [x^{(a+1)} x^{(b)} \ldots x^{(c)} + x^{(a)} x^{(b+1)} \ldots x^{(c)} + x^{(a)} x^{(b)} \ldots x^{(c+1)}] \tag{12}$$

hence if we form h using this expression we see that

$$h(s_1,s_2,\ldots s_r) = \sum_{(a,b,\ldots c)\in I(n-1,r)} \beta_{a,b,\ldots c}(s_1+s_2+\ldots+s_r)(s_1^a s_2^b \ldots s_r^c) \tag{13}$$

Hence the indicated factorization of h is a necessary condition.

To establish sufficiency, suppose that we write $h(s_1,s_2,\ldots s_r) = (s_1+s_2+\ldots s_r) m(s_1,s_2,\ldots s_r)$ with m given in symmetric form by

$$m(s_1,s_2,\ldots s_r) = \sum_{(a,b,\ldots c)\in I(n-1,r)} \gamma_{a,b,\ldots c} s_1^a s_2^b \ldots s_r^c \tag{14}$$

Form the function

$$\tilde\eta = \sum_{(a,b,\ldots c)\in I(n-1,r)} \gamma_{a,b,\ldots c} x^{(a)} x^{(b)} \ldots x^{(c)} \tag{15}$$

It follows by differentiation (as used above) that the derivative of $\tilde\eta$ is actually equal to the integrand and so $\eta = \tilde\eta(a) - \tilde\eta(b)$ and the proof is complete.

The operation of passing from the form to the symmetric function $h(s_1,s_2,\ldots s_r)$ establishes a one to one correspondence between homogeneous forms of degree r in $x, x^{(1)}, \ldots x^{(n)}$ and symmetric polynomials in $s_1, s_2, \ldots s_r$. This operation is vaguely analogous to the assignment of a symbol to an operator in partial differential equations and we will call h the symbol of the form.

Notice that in the case where p = 2 there is a particular simplification which occurs in the path independence condition. In this case we have independence if

$$h(s_1,s_2) = \sum \alpha_{ab} s_1^a s_2^b = (s_1+s_2) m(s_1,s_2) \tag{16}$$

however $h(s_1,s_2)$ contains $s_1+s_2$ as a factor if and only if h vanishes upon setting $s_1 = -s_2$. Thus we see that

$$h(s_1,-s_1) = 0 \tag{17}$$

is equivalent to the condition which appears in [4] and [6]. (Incidentally there are minor defects in all the previously published proofs of the p=2 case that I know of. The present proof plan may be the best way to proceed even if the case p=2 is the only one of interest.)

## 3. Sufficient Conditions for Optimality

As is made clear in [6], all of the conventional least squares theory can be derived from the inequality for real n-vectors $\langle x,x\rangle \geq 0$ via completing the square. In the present setting a key role is played by the inequality in real variables a, b

$$0 \leq a^{2r}+2rab^{2r-1}+(2r-1)b^{2r}$$
$$= (a+b)^2(a^{2r-2}-2a^{2r-3}b+3a^{2r-4}b^2-\ldots+(2r-1)b^{2r-2}) \tag{18}$$

This result is easily verified by differentiating with respect to a and seeing that zero is the minimum value. Moreover $a = -b$ is the only condition under which the inequality is not strict.

The above observation, together with lemma 1 allows us to state what looks a priori to be a rather special sufficient condition for $u=-q(D)y$ to be an optimal control for (3), (4). In actual fact this condition turns out to be much more general than it appears. Here and below we call a control law stabilizing of it makes the null solution asymptotically stable in the large.

Lemma 2: Let u and y be related by (3). A sufficient condition for a stabilizing control law $u = -q(D)y$ to be optimal relative to all other (linear or nonlinear) stabilizing control laws is that

$$L(y,y^{(1)},\ldots,y^{(n-1)})$$
$$= (2r-1)(q(D)y)^{2r-1}(p(D)y-\frac{1}{q_{n-1}}Dq(D)y)$$
$$+2r(q(D)y)^{2r}+\dot{E} \tag{19}$$

where $\dot{E} = \dot{E}(y,y^{(1)},\ldots y^{(n-1)})$ satisfies the condition of lemma 1.

Proof: We write using the hypothesis

$$\eta = \int_0^\infty L(y,y^{(1)},\ldots y^{(n-1)})+(p(D)y)^{2r}dt$$
$$= \int_0^\infty (2r-1)(q(D)y)^{2r-1}(p(D)y-\frac{1}{q_{n-1}}Dq(D)y)$$
$$+2r(q(D)y)^{2r}+(p(D)y)^{2r}dt \ + E\Big|_0^\infty$$
$$= \int_0^\infty (p(D)y+q(D)y)^2\psi(y,y^{(1)},\ldots y^{(n-1)})dt$$
$$-\frac{1}{q_{n-1}}\cdot\frac{2r-1}{2r}(q(D)y)^2\Big|_0^\infty + E\Big|_0^\infty \tag{20}$$

where $\psi$ in the last integral is given by

$$\psi(y,y^{(1)},\ldots y^{(n-1)}) = (p(D)y)^{2r-2}$$
$$-2(p(D)y)^{2r-3}(q(D)y)+\ldots+(2r-1)(q(D)y)^{2r-2} \tag{21}$$

and is nonnegative in view of the first inequality of this section. This representation makes it clear that the way to minimize $\eta$ is to let $p(D)y+q(D)y=0$ because $E(\infty)$ and $E(0)$ are fixed by the initial data and the stabilizing assumption.

## 4. The Main Theorems

We now characterize those L's which have optimal control laws which are linear and at the same time set the stage for the solution of the inverse problem.

Theorem 1: The linear dynamics (3) and the performance measure (4) give rise to an optimal stabilizing control law which is linear if and only if L is homogeneous of degree 2r and

$$L(y,y^{(1)},\ldots y^{(n-1)})$$
$$= (2r)(q(D)y)^{2r}-(2r-1)p(D)y(q(D)y)^{2r-1}+\dot{\mathcal{E}} \tag{22}$$

for some polynomial $q(D)=(q_{n-1}D^{n-1}+q_{n-2}D^{n-2}+\ldots+q_o)$ and some $\mathcal{E} = \mathcal{E}(y,y^{(1)},\ldots,y^{(n-1)})$ such that $p(D)+q(D)$ has all its zeros in Re $s < 0$.

Proof: We have already seen in Lemma 2 that this condition is sufficient for optimality. We establish its necessity here.

Suppose that there is an optimal stabilizing control law $u(t) = h(D)y(t)$. Then since L is a polynomial, the minimum value of $\eta$ corresponding to an initial value of y, $y^{(1)},\ldots,y^{(n-1)}$ will be a polynomial in these variables. (This was worked out in Liapunov's thesis!) Call this polynomial $\phi$. By computing the derivative of $\phi$ along solutions of $p(D)y = -h(D)y$ we get

$$\sum_{i=0}^{n-1}\frac{\partial\phi}{\partial y^{(i)}}y^{(i+1)} = -L(y,y^{(1)},\ldots y^{(n-1)})-(h(D)y)^{2r} \tag{23}$$

provided we set

$$y^{(n)} = -(p_{n-1}y^{(n-1)}+\ldots+p_oy)-(h(D)y) \tag{24}$$

Moreover, and it is surprising how much this point reveals, if we exchange $h(D)y$ for $\alpha h(D)y$ in these equations and replace $\phi$ by $\phi(\alpha,y,y^{(1)},\ldots y^{(n-1)})$ it follows from the optimality of $h(D)y$ that

$$\frac{\partial\tilde{\phi}}{\partial\alpha}\Big|_{\alpha=1} = 0 \tag{25}$$

We use this fact as follows. In (23) and (24) replace $h(D)$ by $\alpha h(D)$ and replace $\phi$ by $\tilde{\phi}$. Then use this new version of (24) to eliminate $y^{(n)}$ in the new version of (23). This yields

$$\sum_{i=0}^{n-2}\frac{\partial\tilde{\phi}}{\partial y^{(i)}}y^{(i+1)}-\frac{\partial\tilde{\phi}}{\partial y^{(n-1)}}[(p_{n-1}y^{(n-1)}+\ldots+p_oy)$$
$$+\alpha^{2r}(h(D)y)] = -L(y,y^{(1)},\ldots,y^{(n-1)})$$
$$-\alpha^{2r}(h(D)y)^{2r} \tag{26}$$

Let $\tilde{\phi}'$ indicate the derivative of $\tilde{\phi}$ with respect to $\alpha$. Then

$$\sum_{i=0}^{n-2}\frac{\partial\tilde{\phi}'}{\partial y^{(i)}}y^{(i+1)}-\frac{\partial\tilde{\phi}'}{\partial y^{(n-1)}}[(p_{n-1}y^{(n-1)}+\ldots p_oy)$$
$$+\alpha^{2r}(h(D)y)]-2r\alpha^{2r-1}\frac{\partial\tilde{\phi}}{\partial y^{(n-1)}}(h(D)y)$$
$$= -2r\alpha^{2r-1}(h(D)y)^{2r} \tag{27}$$

In order for $\frac{\partial\tilde{\phi}}{\partial\alpha}$ to vanish at $\alpha = 1$ it is clear that the last term on the left must equal the term on the right. But at $\alpha = 1$ $\tilde{\phi}$ equals $\phi$ so we obtain the key equation

$$2r \frac{\partial\phi}{\partial y^{(n-1)}} h(D)y = 2r(h(D)y)^{2r} \quad (28)$$

The first consequence we draw from (28) is that since $0 \leq \phi$ cannot depend linearly on $y^{(n-1)}$ it must satisfy

$$\phi = \beta(h(D)y)^{2r} + \phi_o(y, y^{(1)}, \ldots y^{(n-2)}) \quad (29)$$

where $\beta = 1/h_{n-1}$ with $h_{n-1}$ the coefficient of $D^{n-1}$ in $h(D)$ (necessarily nonzero).

Using our new representations for $\phi$ in (23) we get

$$\sum_{i=0}^{n-2} \frac{\partial\phi_o}{\partial y^{(i)}} y^{(i+1)} + 2r\beta(h(D)y)^{2r-1}(Dh(D)y)$$
$$= -L(y, y^{(1)}, \ldots y^{(n-1)}) - (h(D)y)^{2r} \quad (30)$$

with the understanding that the $h_{n-1}y^{(n)}$ terms which appears in the second term on the left must be eliminated using (24). This is conveniently done by subtracting $h_{n-1}(p(D)y+h(D)y$ from $Dh(D)y$ yielding

$$2r(h(D)y)^{2r-1}(h(D)y+p(D)y)$$
$$= -L(y, y^{(1)}, \ldots y^{(n-1)}) - (h(D)y)^{2r} + \dot{E} \quad (31)$$

where

$$\dot{E} = -\frac{d}{dt}(\phi_o + \beta(h(D)y)^{2r}) \quad (32)$$

This then yields the equation of the theorem statement if we let $q(D) = h(D)$.

We still must establish that L is homogeneous of degree 2r. This is actually a consequence of the nonnegativity of L in the following way. Let $L_o$ be either the collection of highest degree terms in L or the collection of lowest degree terms in L. If its degree is not 2r then by (24) we see that $L_o$ does not depend on $y^{(n-1)}$ and hence $L_o$ is exact in the sense of lemma 1. Since it is exact the highest derivative of y present enters linearly and since $L_o$ is of extreme degree in L (either the highest or the lowest) this is incompatible with $L \geq 0$. Hence $L_o$ must be of degree 2r.

Theorem 2: Given the linear system (3) there exists an L in $P^+$ such that $u = -q(D)y$ is optimal for (4) among all stabilizing control laws if and only if

$$\int_a^b [p(D)y + (\tfrac{2r-1}{2r})q(D)y][q(D)y]^{2r-1} dt$$
$$= \phi(x)\Big|_a^b + \int_a^b \psi(x)dt \quad (33)$$

with $\psi$ in $P^+$ and all zeros of $p(D)+q(D)$ in Re $s < 0$.

Proof: Suppose such a decomposition exists. Let $L=\psi$ and apply the inequality of lemma 2. On the other hand, suppose that $-q(D)x = u$ is optimal. Then applying theorem 1 we see that we can take $\psi$ to be $L(y, \dot{y}, \ldots y^{(n-1)})$ and $\phi$ to be $-E$.

Remark: Note that $\phi(x)$ is necessarily nonnegative. In Brockett and Willems [3] the conditions under which a decomposition of the type called for here are investigated extensively and a quite general sufficient condition is developed (see theorems 4 and 5 of Part II). The basic idea is, of course, that the system

$$p(D)y + \frac{2r-1}{2r} q(D)y = u; \quad \tilde{y}(t) = y^{2r-1}$$

should be "passive" or "positive" in a special and strong way, namely, in a way that is revealed by polynomial $\phi$ and $\psi$ functions.

If we give up the requirement on L that it be polynomial, then the class of linear control laws which are optimal becomes wider still. The reason for this is that it is best understood in the context of this theorem and may be explained by saying that there are positive systems which are not polynomially positive. This is brought out forcefully by the results of O'Shea [8]. In view of this one may also state a related result.

Theorem 3: Given the linear system (3) there exists a nonnegative function L such that for (3), (4) the control law $u = -q(D)x$ is optimal among all stabilizing control laws provided

$$\int_{-\infty}^{\infty} (p(D)y + \frac{2r-1}{2r}q(D)y)(q(D)y)^{2k} dt \geq 0 \quad (34)$$

for all paths y of compact support and $p(D)+q(D)$ has all its zeros in Re $s < 0$.

One connection with Fourier methods should be pointed out. One sees easily (for example by letting $y(t)=\sin \omega t$) that a necessary condition for the hypothesis of theorems 2 or 3 hold is that

$$\mathrm{Re}(p(i\omega) + \frac{2r-1}{2r} q(i\omega))q(-i\omega) \geq 0 \quad (35)$$

as a consequence the Nyquist locus of the system $g(s) = q(s)/p(s)$ satisfies

$$\mathrm{Re}[1 + (\tfrac{2r-1}{2r})g(s)]g(-s)\Big|_{s=i\omega} \geq 0 \quad (36)$$

which means $g(i\omega)$ must avoid a circle centered on the negative real axis having a diameter the segment $[-\frac{2r}{2r-1}, 0]$. See figure 1 where the disks for r=1 and r=∞ are sketched.

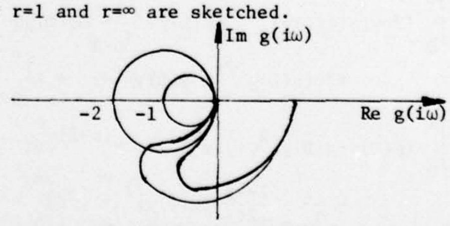



Figure 1: The Forbidden Disks for r=1 and r=∞

It is, however, to be emphasized that for r>1 this Nyquist-like condition is only a necessary condition for optimality. Theorem 2 must be checked if the Nyquist locus passes into the r=1 disk but not into the r=∞ disk.

## 5. Stability

As a direct result of the inequality of Theorem 2 we can establish the following stability property of optimal systems. The special case r=1

is well known.

Theorem 4: If $u = -(q(D)y)$ is an optimal stabilizing control law for system (3) and performance measure (4) then the null solution of the closed loop system obtained by setting $u(t) = -k(t)q(D)y(t)$ will be stable for all k such that

$$\frac{2r-1}{2r} \leq k(t) < \infty \quad (37)$$

Proof: The closed loop equations are

$$p(D)y(t)+k(t)q(D)y(t) = 0 \quad (38)$$

multiply this equation by $(q(D)y)^{2r-1}$ and rearrange it as

$$(q(D)y)^{2r-1}[p(D)y+(\frac{2r-1}{2r})q(D)y] = -(k-\frac{2r-1}{2r})(q(D)y)^{2r} \quad (39)$$

using theorem 2 we see that

$$\frac{d}{dt}\phi = -\psi-(k-\frac{2r-1}{2r})(q(D)y)^2 \quad (40)$$

Thus $\phi$ is a Liapunov function whose derivative is negative semidefinite provided the hypothesis holds.

## 6. An Example

To illustrate some of these ideas we consider the second order system

$$\ddot{y} = u; \quad \eta = \int_0^\infty L(y,\dot{y})+u^{2r}dt \quad (41)$$

According to theorem 1 we must have for linear optimality

$$L(y,\dot{y}) = 2r(\alpha\dot{y}+\beta y)^{2r}-(2r-1)(\alpha\dot{y}+\beta y)^{2r-1}(\frac{\beta}{\alpha})\dot{y}+\gamma\dot{y}y^{2r-1} \quad (42)$$

because an exact form is of the type $\gamma\dot{y}y^{2r-1}$. To find out if L is nonnegative we introduce $\dot{y}/y=z$. The condition for nonnegativity is then

$$2r(\alpha z+\beta)^{2r}-\frac{(2r-1)\beta}{\alpha}z(\alpha z+\beta)^{2r-1}+\gamma z \geq 0. \quad (43)$$

A lengthy calculation shows that this is nonnegative for some $\gamma$ if and only if the obvious necessary condition holds on the highest power of z,

$$2r\alpha^{2r} - \frac{(2r-1)\beta}{\alpha}\alpha^{2r-1} \geq 0 \quad (44)$$

This is equivalent to

$$\alpha^2/\beta \geq \frac{2r-1}{2r} \quad (45)$$

Notice that for r=1 optimal systems have damping constant larger than .7071 but for $r = \infty$ it need only be larger than 1/2.

## 7. The Riccati-Like Formulation

For the sake of completeness we give here the "state space" version of the main conclusions. To begin with we remind the reader of some notation which will play a role here as it has in some other recent work in system theory. If x is an n-tuple, $(x_1,x_2,\ldots,x_n)$ then by $x^{[r]}$ we mean the $\binom{n+r-1}{r}$-tuple

$$x^{[r]} = \begin{bmatrix} x_1^r \\ \alpha_1 x_1^{r-1}x_2 \\ \alpha_2 x_1^{r-1}x_3 \\ \vdots \\ x_n^r \end{bmatrix} \quad (46)$$

where the $\alpha_i$ are chosen in such a way as to have

$$\langle x^{[r]},x^{[r]}\rangle = ||x||^{2r}$$

If y = Ax then there exists $A^{[r]}$ such that $y^{[r]}=A^{[r]}x^{[r]}$. If $\dot{x} = Ax$ there exists $A_{[r]}$ such that

$$\frac{d}{dt}x^{[r]} = A_{[r]}x^{[r]} \quad (47)$$

See our earlier papers [9,10] for some properties of "super-[r]" and "sub-[r]" and some control theoretic applications.

Here we are especially concerned with quadratic forms in $x^{[r]}$. It is clear that for each symmetric matrix Q (we write ${}^ty$ for transpose)

$$\phi(x) = {}^tx^{[r]}Qx^{[r]} \quad (48)$$

is a form homogeneous of degree 2r. If r and di x exceed 1 there are nonzero symmetric Q's such that ${}^tx^{[r]}Qx^{[r]} = 0$ so we see that the representations of $\phi$ in this way is not unique. However the nonuniqueness is easily described. If $Q_1$ and $Q_2$ define the same homogeneous form then $Q_1-Q_2$ is equivalent to zero. Thus we see that given n and r there is a subspace, K(n,r), of the space of all symmetric $\binom{n+r-1}{r}$-order matrices such that any two such matrices define the same homogeneous form if and only if their difference is in K(n,r).

The following lemma should be borne in mind when interpreting the equations which follow.

Lemma 1: The space K(n,r) is invariant for operators of the form

$$L(Q) = QA_{[r]}+{}^tA_{[r]}Q \quad (49)$$

Moreover there exists a complementary space which is also invariant.

Proof: If $\phi(x) = {}^tx^{[r]}Qx^{[r]}$ then along solutions of $\dot{x} = Ax$ we see that

$$\frac{d}{dt}{}^tx^{[r]}Qx^{[r]} = {}^tx^{[r]}L(Q)x^{[r]} \quad (50)$$

but if Q is in K(n,r) then the derivative is clearly zero and hence L(Q) is in K(n,r).

If we regard the space of symmetric matrices of order $\binom{n+r-1}{r}$ as an inner produce space with $\langle M,N\rangle = \mathrm{tr}\, M^tN^r$ then it makes sense to ask if the orthogonal complement of K(n,r) is also invariant. To see that it is, let K belong to K(n,r). Then for all Q

$$\mathrm{tr}[K({}^tA_{[r]}Q+QA_{[r]})]=\mathrm{tr}[Q(K{}^tA_{[r]}+A_{[r]}K)] = 0 \quad (51)$$

this last fact coming from the fact that ${}^tA_{[r]} = ({}^tA)_{[r]}$. (See [10].)

Now consider the linear system (1) with the performance measure

$$\eta = \int_0^\infty {}^t x^{[r]} M x^{[r]} + u^{2r}\, dt \tag{52}$$

In terms of this notation equation (23) becomes

$$Q(A+bc)_{[r]} + {}^t(A+bc)_{[r]} Q = -M - {}^t c^{[r]} c^{[r]} \tag{53}$$

and (28) becomes

$$Q(bc)_{[r]} + {}^t(bc)_{[r]} Q = -2r\, {}^t c^{[r]} c^{[r]} \tag{54}$$

In the case r=1 this pair of equations specializes to the familiar pair from least squares theory,

$$QA + {}^tAQ - Qb\,{}^tbQ = -M; \quad Qb = {}^tc \tag{55}$$

Assuming controllability of (A,b), according to theorem 1 these equations have a solution (Q,c) if and only if for some ${}^tc \in \mathbb{R}^n$ with $cb \neq 0$ we have

$${}^tx^{[r]} M x^{[r]} = 2r(cx)^{2r} - (cb)^{-1}(2r-1)(cx)^{2r-1}(cAx) + \dot{E} \tag{56}$$

where $\dot{E} = \dot{E}(y, y^{(1)}, \ldots, y^{(n-1)})$ is a perfect differential. That is

$$\dot{E} = {}^tx^{[r]}(RA_{[k]} + {}^tA_{[k]}R)x^{[r]} \tag{57}$$

for some $R = {}^tR$ having the property that for all ${}^td$ in $\mathbb{R}^n$

$$F(d) = R(bd)_{[k]} + {}^t(bd)_{[k]}R = 0 \tag{58}$$

(This last condition insures that $\dot{E}$ does not depend on $y^{(n)}$.)

The verification of these remarks is left to the reader.

References


1. R.E. Kalman, "When is a Linear Control System Optimal?" Trans. ASME (J. Basic Engrg.) pt. D., March 1964, pp. 51-60.

2. R.W. Brockett. "On the Stability of Nonlinear Feedback Systems," IEEE Applications and Industry, Vol. 83, pp. 443-448, November 1964.

3. R.W. Brockett and J.L. Willems, "Frequency Domain Stability Criteria, Parts I and II," IEEE Trans. on Automatic Control, Vol. 10, pp. 255-261 and 407-413, July and Oct. 1965.

4. R.W. Brockett, "Path Integrals, Liapunov Functions, and Quadratic Minimization," Proc. of the 4th (1966) Allerton Conference, Univ. of Illinois, Urbana, Illinois, pp. 685-697.

5. R.W. Brockett and H.B. Lee, "Frequency-Domain Instability Criteria for Time-Varying and Nonlinear Systems," Proc. of the IEEE, Vol. 55, No. 5, May 1967, pp. 604-619.

6. R.W. Brockett, Finite Dimensional Linear Systems, J. Wiley, N.Y., 1970.

7. J.L. Willems, Stability Theory of Dynamical Systems, Nelsons Ltd., London, 1970.

8. R.P. O'Shea, "A Combined Frequency-Time Domain Stability Criterion for Autonomous Continuous Systems," IEEE Trans. Automatic Control, Vol. AC11, pp. 477-484, July 1966.

9. R.W. Brockett, "Lie Theory and Control Systems Defined on Spheres", SIAM J. on App. Math., Vol. 25, No. 2, Sept. 1973, pp. 213-225.

10. R.W. Brockett, "Lie Algebras and Lie Groups in Control Theory," in Geometric Methods in Control Theory, (D.Q. Mayne and R.W. Brockett, eds.) Reidel, Dordretch, 1973, pp. 43-82.